

AWS NIFS-LHDオープンデータを活用した データ転送の評価

国立情報学研究所
山中顕次郎

AWS NIFS-LHDオープンデータとは

- NIFS LHDの全計測データをAWSのオープンデータスポンサーシッププログラムに基づき、無償公開（AWSのアカウントが無くてもダウンロード可能）した物。
- AWSオブジェクトストレージサービス（Simple Storage Service: S3）にファイル数：約18百万、ファイル容量合計 約2PiBの巨大なデータ群を保管

AWS NIFS-LHD収容ファイルの内訳

C	年	数	容量	C	年	数	容量
1	1998	10,949	7.2 GiB	14	2010	748,787	35.0 TiB
2	1998	42,063	41.4 GiB	15	2011	888,271	43.1 TiB
3	1999	147,463	269.7 GiB	16	2012	804,657	40.7 TiB
4	2000	168,368	454.7 GiB	17	2013	1,278,307	62.7 TiB
5	2001	213,867	644.8 GiB	18	2014	937,069	46.6 TiB
6	2002	153,470	575.9 GiB	19	2017	1,669,605	136.3 TiB
7	2003	222,538	1.0 TiB	20	2018	1,234,487	93.7 TiB
8	2004	367,879	3.4 TiB	21	2019	1,367,519	97.2 TiB
9	2005	612,506	6.7 TiB	22	2020	1,457,055	249.4 TiB
10	2006	143,156	1.9 TiB	23	2021	1,670,475	284.6 TiB
11	2007	711,431	16.7 TiB	24	2022	1,393,957	248.8 TiB
12	2008	584,386	12.6 TiB	25	2024	948,115	208.8 TiB
13	2009	579,008	17.3 TiB	26	2025		

NIFS Large Helical Device (LHD) Experiment

[analytics](#) [diagnostic](#) [archive](#) [computational tomography](#) [diagnostics](#) [digital assets](#) [electrocity](#) [energy](#) [fluid dynamics](#) [image processing](#) [physics](#)
[post-processing](#) [reactions](#) [signal processing](#) [seismology](#) [turbulence](#) [video](#) [xray](#) [xray tomography](#)

Description
 The Large Helical Device (LHD), owned and operated by the National Institute for Fusion Science (NIFS), is one of the world's largest plasma confinement device which employs a heliotron magnetic configuration generated by the superconducting coils. The objectives are to conduct academic research on the confinement of steady-state, high-temperature, high-density plasmas, core plasma physics, and fusion reactor engineering, which are necessary to develop future fusion reactors. All the archived data of the LHD plasma diagnostics are available since the beginning of the LHD experiment, started on 31st of March, 1998.

Update Frequency
 Archived data files are updated nightly when new or revised data are generated in LHD experiment.

License
 This data is available for anyone to use under the "Rights and Terms"

Documentation
https://www.lhd-nifs.ac.jp/pub/Repository_en.html

Managed By
 NIFS
 See all datasets managed by NIFS.

Contact
 For any questions regarding data delivery or any general questions regarding the LHD Experiment data repository, please send email to the Data Acquisition and Analysis group at Comp_DAE@nifs.ac.jp.

How to Cite
 NIFS Large Helical Device (LHD) Experiment was accessed on DATE from <https://registry.opendata.aws/nifs-lhd/>.

Usage Examples

Tutorials

- Data handling in the LHD experiments -- Guide for the collaborator (pdf) by Satoshi Ohdachi

Tools & Applications

- How to retrieve diagnostic raw data and/or primarily processed data using "Retrieve" by NIFS LABCOM

Publications

- Data Acquisition and Management System of LHD by Hideya Nakajima et al.

Resources on AWS

Description
 LHD Diagnostic data

Resource type
 S3 Bucket

Amazon Resource Name (ARN)
[arn:aws:s3:::nifs-lhd](#)

AWS Region
[ap-northeast-1](#)

AWS CLI Access (No AWS account required)
[aws s3 ls --no-sign-request s3://nifs-lhd/](#)

Explore
[Browse Bucket](#)

<https://registry.opendata.aws/nifs-lhd/>

概要

- AWS NIFS-LHDデータの効率的な利用方法を探す
- 利用方法
 - データの取得方法・移動方法にフォーカス
 - データを使った解析は対象外
- 調査項目
 1. S3操作ツールによる性能の違い
 2. 取得データを遠隔利用する際の効率的な転送方法

1. S3操作ツールによる性能の違い (1)

- S3を操作するツールは、純正のaws s3コマンドの他に、様々な物が存在。
 - Linode CLI, s3cmd, s4cmd, s5cmd等
 - 事前調査[1]によるとs5cmdが速い
 - aws s3コマンドとs5cmdを比較

テスト2 - 1MBのファイル100個転送した時間

ツール	秒数
Linode CLI	2分28秒
s3cmd	1分26秒
s4cmd	18秒
s5cmd	14秒

[1] アカマイ・テクノロジーズ合同会社、「CLIを利用して Linode Object Storage にファイルを最速でアップロードする」、
<https://qiita.com/iss802/items/3a4b9865c6f88373e437>

1. S3操作ツールによる性能の違い (2)

● テスト条件：

- AWS東京リージョンのEC2サーバからaws s3とs5cmdを使って各種操作を実施。
- ツールはチューニングは行なわず、標準設定を使用
- OS：Amazon Linux 2023
- データ：2003年第7サイクルの全データ、222,538ファイル、約1,141GB (1.04 TiB)

● 結果：s5cmdが桁違いに高速

ツール	ダウンロード1	アップロード	コピー	ダウンロード2	移動(mv)	消去(rm)
aws s3	56m44.580s 2.68 Gbps	81m6.616s 1.88 Gbps	103m34.876s 1.47 Gps	- -	100m36.203s 1.51 Gbps	10m34.112s 14.40 Gbps
s5cmd	8m37.477s 17.64 Gbps	8m32.681s 17.81 Gbps	3m1.099s 50.41 Gbps	2m48.225s 54.27 Gbps	4m6.984s 40.39 Gbps	3m46.028s 36.96 Gbps
instance	m6i.12xlarge, i4i.8xlarge (NW帯域:18.75 Gbps)			i4i.32xlarge (NW帯域:75 Gbps)		

NW帯域がボトルネックと判明したので、より高性能なサーバ（インスタンス）を使用

1. S3操作ツールによる性能の違い (3)

- s5cmdの限界：
 - 大きな1つのファイルのアップロード・ダウンロードでは、aws s3とあまり速度が変わらない。 → 多数のファイルを扱う時に、高速性が発揮される。
 - 通常のOS経由のIOを使う為、高速SSDの性能を活かせない可能性がある。 → SSDの性能限界を引き出すには、非同期IOが必要なのだが。 . . .
 - マルチコアによる並列処理で高速化を図っているのでCPU数が少ないサーバだと、高速性を発揮できない可能性がある。
 - 通信距離が長くなる程、性能が低下する。 → 調査項目2のテーマ

(参考) s5cmdについて

- <https://github.com/peak/s5cmd> で公開。 MIT license
- pipでインストール可能 (但し、python 3.11以上が必要)

2. 取得データを遠隔利用する際の効率的な転送方法（1）

- s5cmdをもってしても、通信距離が長くなる程、転送速度が低下する（下表）

AWS NIFS-LHDのダウンロード時間（データは項目1と同じで約1TiB）

From	推定RTT(*)	転送時間	転送速度	サーバ仕様
東京リージョン	< 1 ms	2min 48.2s	54.27 Gbps	i4i.32xlarge (NW帯域:75 Gbps)
北海道北見	20 ms	5min 7.5s	29.69 Gbps	SINET6 データ転送実験用サーバ (NW帯域: 100Gbps x2)
パリリージョン	250 ms	17min 7.3s	8.89 Gbps	i4i.32xlarge (NW帯域:75 Gbps)

(*) s3サーバにpingする事は出来ないため、推定値

- NIIのファイル転送ツールMMCFTPと組み合わせることで、高速化できるのでは？
 1. データは、東京リージョンで取得する（50 Gbps超の速度入手可能）
 2. 取得したデータをMMCFTPで遠隔地（パリリージョン）に送る
 3. トータルの転送時間は、どうなるか？

2. 取得データを遠隔利用する際の効率的な転送方法 (2)

- テスト条件：

- EC2サーバは東京・パリ共にi4i.32xlarge (NW帯域:75 Gbps)を利用。
- MMCFTPでの転送では、新たに開発したボリューム転送機能^(*1)を利用。
- 通信経路は、インターネット経由と仮想内部ネット直接通信 (VPCピアリング) の2種類を試行

(*1) ディスク全体またはそのパーティションを丸ごと転送する機能。多数のファイルが含まれていても、一つのブロックファイルとして扱う為、MMCFTPで高速転送できる。受信結果はDense RAWのイメージファイルとなり、ループバックマウントにより通常のファイルシステムとしてアクセスできる。

結果： MMCFTPを併用した方が、s5cmd直接より、高速転送可能。

s5cmd	17m7.335s / 8.89 Gbps			
	ダウンロード (東京)	umount	MMCFTP転送 ^(*2)	トータル
インターネット	2m 48.225s / 54.27 Gbps	0m 58.091s	4m 53s / 32.69 Gbps	8m 39s / 17.58 Gbps
VPCピアリング			3m 18s / 48.29 Gbps	7m 4s / 21.5 Gbps

(*2) MMCFTPの転送速度として、インターネット経由では約33.6Gbpsを指定し、VPCピアリングでは約50Gbpsを指定した。AWSでは、インターネット経由通信はNW帯域が表定の半分に制限される。VPCピアリングではその制限は無いが表定帯域が保証される訳ではない。

実験結果のまとめとサジェスション（1）

- AWS S3を操作するツールとしては、s5cmdを推奨.
- AWS NIFS-LHDデータを最高性能で利用するには？
 - AWS 東京リージョンのEC2サーバを使うのが、确实.
 - 50 Gbps超のNW性能・ストレージ性能を持つオンプレミスサーバは、高価.
 - 高価なサーバを用意しても、インターネット経由では組織のNW管理部門が提供するファイアウォールの性能がボトルネックになる事が多い。（但し、SINETクラウド接続サービスを使って、L2VPN接続によりファイアウォール無しでNIFS-LHDデータを使う方法は有り得る）
- EC2サーバでも、高性能サーバは高価.
 - 通常は低性能サーバを使い、必要な時だけ高性能サーバに切り替える（インスタンスタイプ変更する）のが良い.

実施結果のまとめとサジェスション（2）

- データディスクの選択にも注意が必要
 - 実験では、当初 EBS（永続性のあるNWディスク）を使ったが、非常にコストが掛かった為、インスタンスストア（内蔵SSD）を持つインスタンスに切り替えた.
 - インスタンスストアは、サーバを停止すると内容が消去されるが、ストレージ読み書きで課金されないし高性能なので、一時データ置き場として有用.
- MMCFTP併用で遠隔地でも転送速度を上げられるが、致命的な問題「データ転送料が課金される」が有る.
- OSとして、AWS純正だからとの理由で、Amazon Linux 2023を使用したがる、使いたいパッケージが用意されていない事があった。追加料金不要のLinuxは他にも有るので、使い慣れた物を使用するのが良い.

謝辞

- 本研究は以下の研究助成を受けた
 - QST 原型炉開発共同研究
 - NIFS 一般共同研究 (プラズマシミュレータ共同研究) NIFS22KISHI001